

2. Brundage JF, Shanks GD. Deaths from bacterial pneumonia during 1918–19 influenza pandemic. *Emerg Infect Dis* 2008;14:1193–9.

3. Update: novel influenza A (H1N1) virus infections — worldwide, May 6, 2009. *MMWR Morb Mortal Wkly Rep* 2009;58:453–8.

4. Andreasen V, Viboud C, Simonsen L. Epidemiologic characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. *J Infect Dis* 2008;197:270–8.

5. Viboud C, Grais RF, Lafont BA, Miller MA, Simonsen L. Multinational impact of the 1968 Hong Kong influenza pandemic: evidence for a smoldering pandemic. *J Infect Dis* 2005;192:233–48.

Copyright © 2009 Massachusetts Medical Society.

Geographic Dependence, Surveillance, and Origins of the 2009 Influenza A (H1N1) Virus

Vladimir Trifonov, Ph.D., Hossein Khiabanian, Ph.D., and Raul Rabadan, Ph.D.

In April 2009, a new strain of human H1N1 influenza A virus was identified in Mexico. According to the World Health

Organization (www.who.int/csr/don/2009_05_25), as of May 25, 2009, the virus had spread to 43 countries, with 12,515 reported cases and 91 associated deaths, and it has been assessed as having pandemic potential.¹

Genomic analysis of the 2009 influenza A (H1N1) virus in humans indicates that it is closely related to common reassortant swine influenza A viruses isolated in North America, Europe, and Asia (Fig. 1).^{2–4} The segments coding for the polymerase complex, hemagglutinin, nuclear protein, and nonstructural proteins show high similarity with the swine H1N2 influenza A viruses isolated in North America in the late 1990s (Table 1). H1N2 and other subtypes are descendants of the triple-reassortant swine H3N2 viruses isolated in North America. They have spread in swine hosts around the globe and have been found to infect humans.⁵ The segments coding for the neuraminidase and the matrix proteins of the new human H1N1 virus are, however, distantly related to swine viruses isolated in Europe in the early 1990s (Table 2). In particular, the closest isolated relatives of the neuraminidase segment have 94.4% similarity at the nucleotide level with European swine influenza A virus strains from 1992.

In the past few years, there

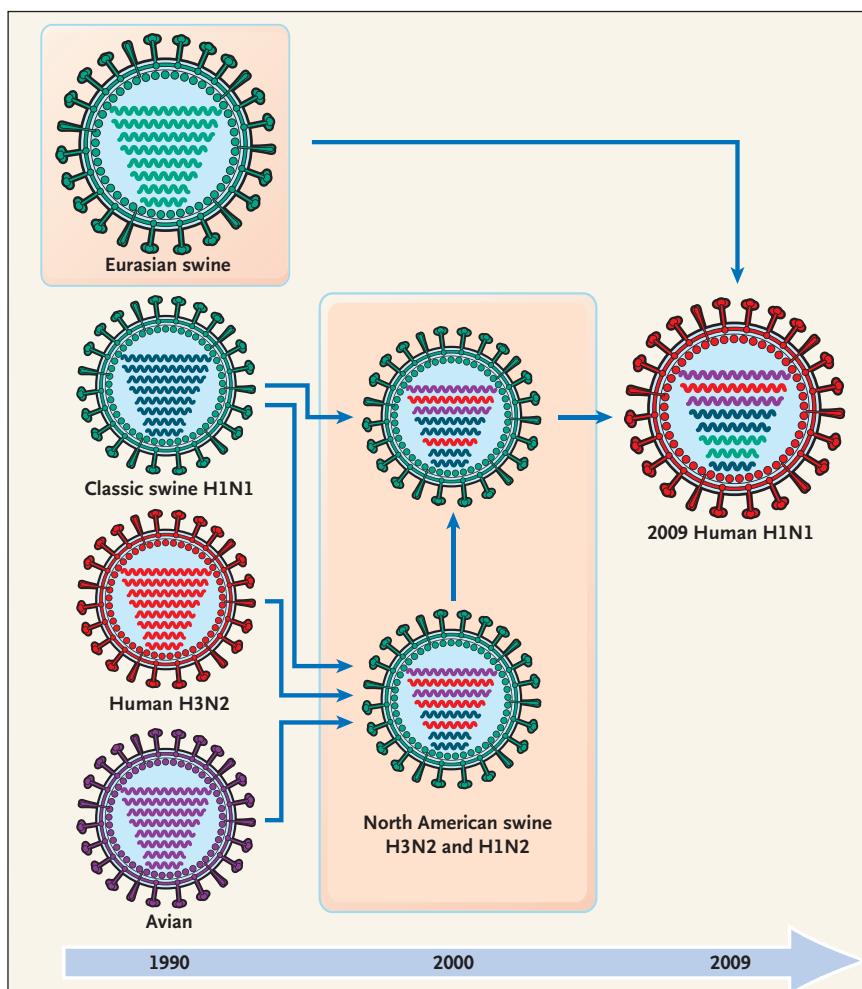


Figure 1. History of Reassortment Events in the Evolution of the 2009 Influenza A (H1N1) Virus.

The eight segments shown within each virus code for the following proteins of the influenza A virus (top to bottom): polymerase PB2, polymerase PB1, polymerase PA, hemagglutinin, nuclear protein, neuraminidase, matrix proteins, and nonstructural proteins. The segments of the human 2009 influenza A (H1N1) virus have coexisted in swine influenza A virus strains for more than 10 years. The ancestors of neuraminidase have not been observed for almost 20 years. The mixing vessel for the current reassortment is likely to be a swine host but remains unknown.

Table 1. Nucleotide Identities of Swine Influenza A Viruses Most Similar to the Ancestor of Segments 1, 2, 3, 4, 5, and 8 of the 2009 Swine-Origin Human Influenza A (H1N1) Virus.*

Virus	Nucleotide Identity (%)						
	PB2	PB1	PA	HA	NP	M1	NS1
North American							
A/Swine/Illinois/100084/2001 (H1N2)	96.1	96.3	96.3	94.9	96.3	88.0	95.4
A/Swine/Indiana/9K035/1999 (H1N2)	96.3	96.6	95.5	95.7	96.6	88.2	95.8
A/Swine/Iowa/930/2001 (H1N2)	96.0	96.3	95.5	92.4	95.7	87.8	95.8
A/Swine/Minnesota/55551/2000 (H1N2)	96.6	96.2	96.2	91.5	96.4	88.1	95.8
A/Swine/North Carolina/98225/2001 (H1N2)	96.6	96.1	95.9	91.6	95.8	88.1	95.9
A/Swine/Ohio/891/2001 (H1N2)	96.1	96.3	95.5	95.2	96.6	88.0	95.6
Asian							
A/Swine/Korea/ASAN04/2006 (H1N2)	95.3	95.4	95.5	93.4	94.9	88.0	94.5
A/Swine/Korea/PZ7/2006 (H1N2)	95.3	95.4	95.6	93.4	95.2	87.9	94.7
A/Swine/Shanghai/1/2007 (H1N2)	95.2	94.8	95.1	90.5	96.2	87.9	94.9

* Strain A/Mexico/InDRE4487/2009 (H1N1) was the reference strain of human 2009 influenza A (H1N1) virus used in the analysis. Alignment data were obtained with the use of the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST), with default settings. Data for neuraminidase are missing because NCBI BLAST did not yield significant alignment, because of low similarity. The recent isolation times of the Asian strains indicate that they are not direct evolutionary precursors of the new strain but instead are recent descendants of a common ancestor. HA denotes hemagglutinin, M1 matrix protein 1, NP nuclear protein, NS1 nonstructural protein 1, PA polymerase PA, PB1 polymerase PB1, and PB2 polymerase PB2.

Table 2. Nucleotide Identities of Swine Influenza A Viruses Most Similar to the Ancestor of Segments 6 and 7 of the 2009 Swine-Origin Human Influenza A (H1N1) Virus.*

Virus	Nucleotide Identity (%)						
	PB2	PB1	PA	HA	NP	NA	M1
European							
A/Swine/Belgium/WVL5/1989 (H1N1)	85.0	87.3	87.4		85.8	92.9	95.4
A/Swine/Denmark/WVL9/1993 (H1N1)	84.8	87.6	87.1		85.4	93.7	96.4
A/Swine/England/WVL7/1992 (H1N1)	84.8	87.2	87.4		85.6	94.4	96.0
A/Swine/France/WVL4/1985 (H1N1)	85.1	87.6	87.6		85.8	93.0	95.3
A/Swine/Spain/WVL6/1991 (H1N1)	84.8	87.5	87.2		85.4	94.3	96.3
North American							
A/Swine/Virginia/670/1987 (H1N1)	85.0	87.2	87.5		85.9	92.2	95.7
Asian							
A/Swine/Hong Kong/5190/1999 (H3N2)			86.6		85.2		97.3
A/Swine/Chachoengsao/NIAH587/2005 (H1N1)	84.5	86.4	86.3	86.5	85.2	91.1	95.1
A/Swine/Chonburi/NIAH589/2005 (H1N1)	84.5	86.4	86.3	86.6	85.2	91.0	95.1
A/Swine/Zhejiang/1/2007 (H1N1)	84.9	86.7	85.7		84.6	91.8	95.0

* Strain A/Mexico/InDRE4487/2009 (H1N1) was the reference strain of human 2009 influenza A (H1N1) virus used in the analysis. Alignment data were obtained with the use of the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST), with default settings. Some data are missing because either the sequence deposited at NCBI is incomplete or because NCBI BLAST did not yield a significant alignment, because of low similarity. The recent isolation times of the Asian strains except A/Swine/Hong Kong/5190/1999 (H3N2) indicate that they are not direct evolutionary precursors of the new strain but instead are recent descendants of a common ancestor. HA denotes hemagglutinin, M1 matrix protein 1, NA neuraminidase, NP nuclear protein, NS1 nonstructural protein 1, PA polymerase PA, PB1 polymerase PB1, and PB2 polymerase PB2.

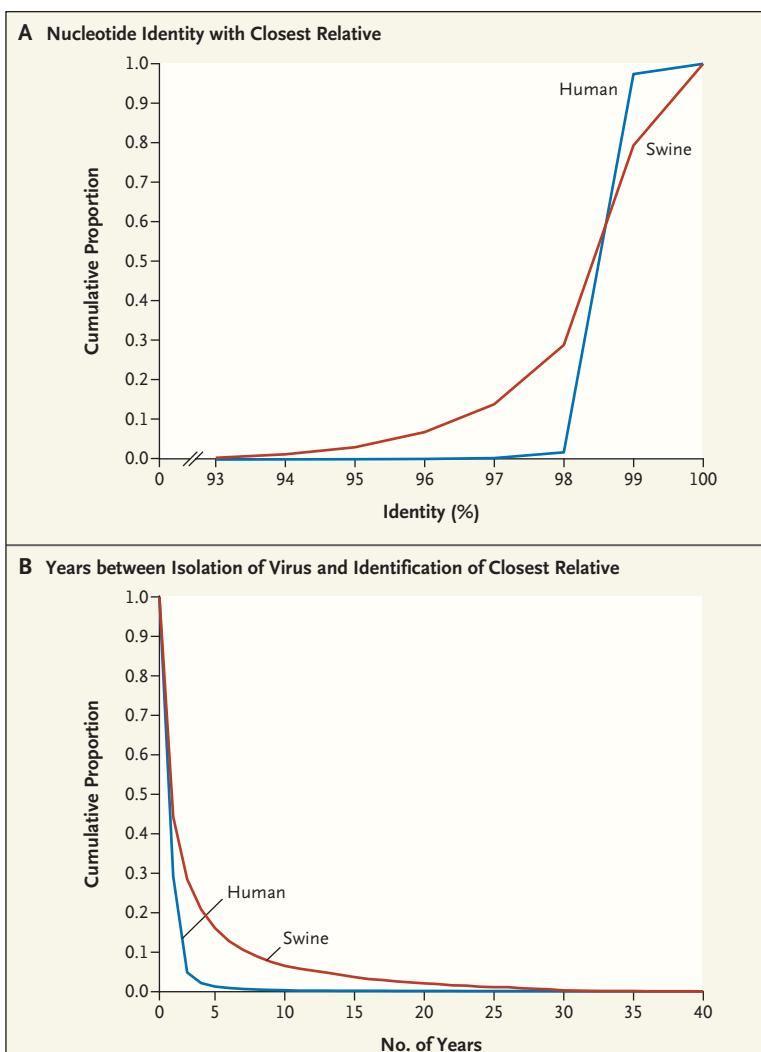


Figure 2. Nucleotide Identities and Numbers of Years between Initial Isolations of Influenza A Viruses and Their Closest Relatives Deposited in the NCBI Database.

Panel A shows the cumulative proportion of sequences with a given degree of nucleotide identity with their closest relative. Panel B shows the cumulative proportion of sequences appearing a given number of years after the closest relative was identified.

has been a worldwide effort to isolate and sequence the genomes of influenza A viruses, which has led to the depositing of more than 46,000 sequences in the Influenza Virus Resource of the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html).

As of May 25, 2009, the NCBI database included sequences from more than 220 strains from the 2009 swine-origin human influenza A (H1N1) virus isolated at various sites around the world. Consequently, the origin and recent history of new strains can be inferred from

study of the most similar deposited sequences. The percentage of matching nucleotides (the nucleotide identity) after nucleotide alignment, as determined with the use of the NCBI Basic Local Alignment Search Tool (BLAST) or other tools, is a common measure of similarity used by researchers in the field.

Figure 2 shows the nucleotide identities and the numbers of years between the initial isolation of a given sequence of influenza A virus deposited in the NCBI database and the initial isolation of its closest relative. A total of 98% of all sequences of human influenza A viruses have relatives with at least 99% nucleotide identity, and 95% have a relative that was initially isolated within 2 years before their own first appearance. These numbers suggest that researchers have sampled human influenza A viruses efficiently — and point to a high degree of homogeneity among human influenza A viruses. Swine influenza A viruses have not been sampled as efficiently as human influenza A viruses; nevertheless, 86% of all segments from such strains have relatives with at least 99% nucleotide identity, and 71% have a relative that was initially isolated within 2 years before their own first appearance. Only 2% of all swine influenza A virus sequences have 94% nucleotide identity with their closest relative, and in 2% of cases, the closest relative appeared at least 20 years earlier.

Figure 3 shows the numbers of sequences from human and swine hosts isolated on various continents and deposited in the

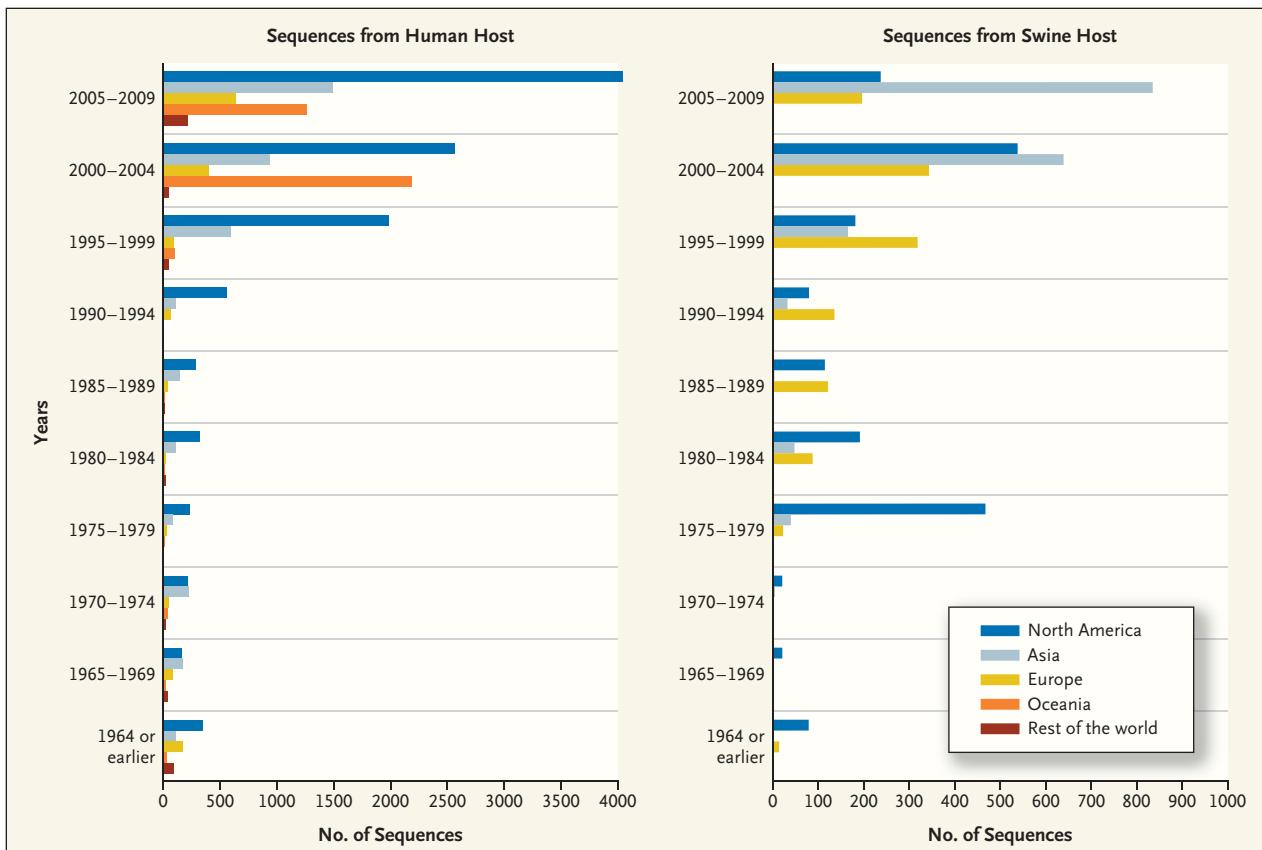


Figure 3. Numbers of Sequences of Influenza A Viruses from Human and Swine Hosts Isolated on Various Continents and Deposited in the NCBI Database.

Sampling has increased in recent years. No sequences of swine influenza A virus strains from Africa, Oceania, or South America have been deposited in the database.

NCBI database. The numbers have increased dramatically over the past few years, and geography plays an important role. Most isolates of human influenza A viruses are from North America, Oceania, Asia, and Europe, and though there are many swine influenza A viruses from North America, Asia, and Europe, there are none from Africa, Oceania, or South America. North American and European swine influenza A (H1N1) viruses show strong geographic homogeneity, whereas some Asian

isolates contain an admixture of both North American and European lineages (Fig. 4). Although human influenza A viruses travel around the world with their hosts, swine viruses on different continents have largely distinct lineages.

Given both the dependence of the distribution of swine influenza A viruses on geographic location and the lack of sampling in certain parts of the world, it is perhaps not surprising that the ancestors of the new human influenza A (H1N1)

virus have gone unnoticed for almost two decades. Only more efficient surveillance could prevent such an event from happening in the future.

No potential conflict of interest relevant to this article has been reported.

This article (10.1056/NEJMp0904572) was published on May 27, 2009, at NEJM.org.

From the Department of Biomedical Informatics, Center for Computational Biology and Bioinformatics, Columbia University College of Physicians and Surgeons, New York.

1. Fraser C, Donnelly CA, Cauchemez S, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009 May 14 (Epub ahead of print).

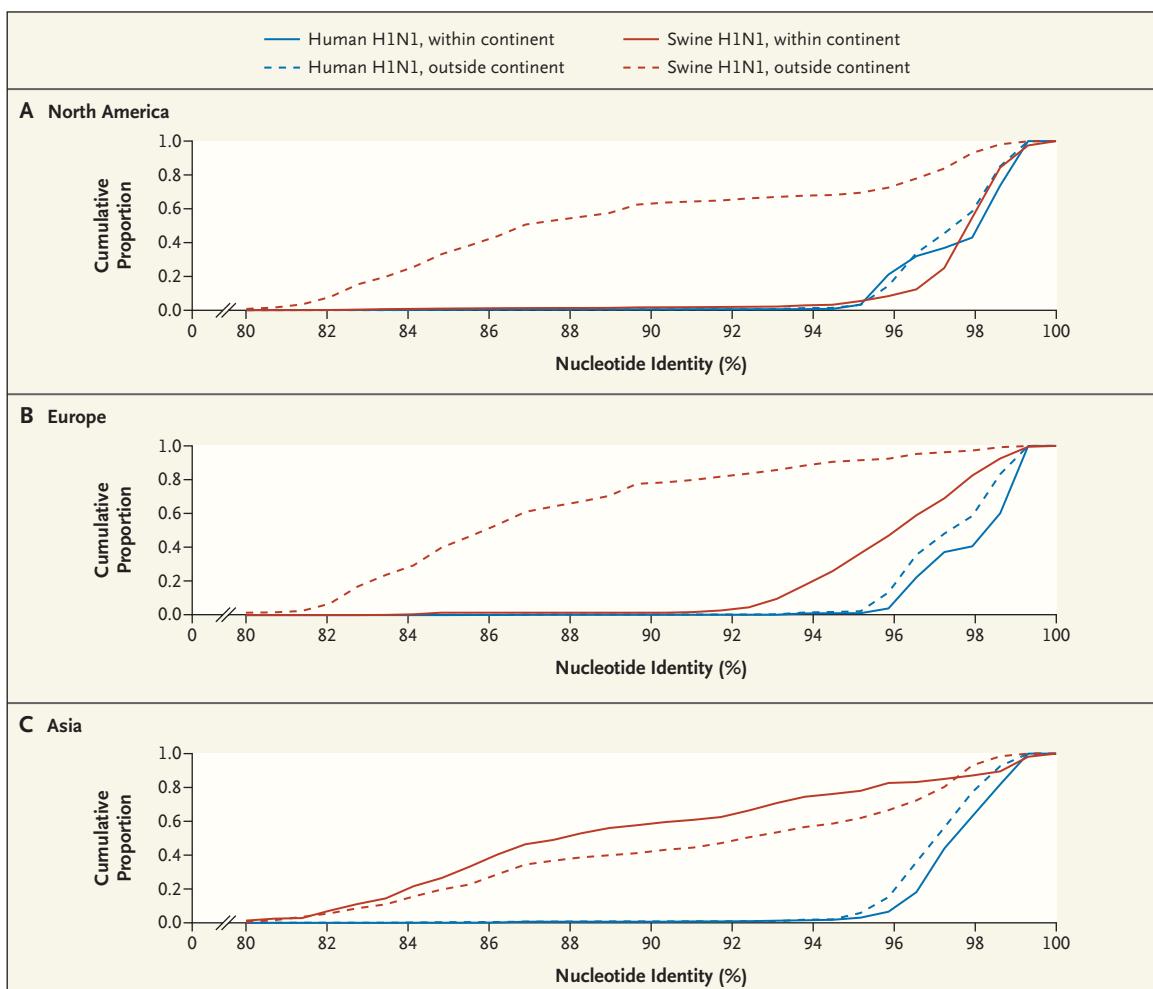


Figure 4. Nucleotide Identities of Human and Swine Influenza A (H1N1) Viruses from within and outside North America, Europe, and Asia.

Human influenza A virus strains show strong worldwide homogeneity. North American and European swine influenza A virus strains are homogeneous within each continent. Asian swine influenza A virus strains contain an admixture of North American and European strains.

2. Trifonov V, Khiabanian H, Greenbaum B, Rabadian R. The origin of the recent swine influenza A(H1N1) virus infecting humans. *Euro Surveill* 2009;14(17):pii=19193.
3. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus

in humans. *N Engl J Med* 2009;360:2605-15.

4. Garten RJ, Davis CD, Russell CA, et al. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 2009 May 22 (Epub ahead of print).

5. Shinde V, Bridges CB, Uyeki TM, et al. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. *N Engl J Med* 2009;360:2616-25.

Copyright © 2009 Massachusetts Medical Society.